

# 上证行业指数收益率的影响因素

## 基于三因素模型的分位数回归分析

方丽婷,李坤明

(厦门大学 统计系,福建 厦门 361005)

**摘要** :Fama and French(1993)提出的三因素模型对许多国家的股市收益率有良好的解释力并得到广泛应用。我们在三因素模型的基础上,利用分位数回归技术全面考察三因素对上证行业指数收益率的影响。实证结果表明:(1)在各个分位点上,市场因子对五类行业指数均有显著的正向影响;(2)规模因子对五个行业指数的影响方向不同,其中对综合指数和地产指数具有显著的负向影响,而对其他三类指数具有显著的正向影响;(3)账面市值因子对五个行业指数的影响方向也不同,其中对综合指数和公用指数具有负向影响,对工业指数和商业指数具有正向影响,对地产指数既有正向影响又有负向影响,同时在部分分位点上,账面市值因子对五个行业指数的影响不显著。

**关键词** 股票市场;行业指数;回报率;Fama-French三因素模型;分位数回归

**文章编号** :1003-4625(2012)10-0008-04 **中图分类号** :F830.91 **文献标志码** :A

### 一、引言

证券资产定价问题始终是金融领域的核心问题。目前国内已有大量文献对股票收益率的决定机制进行分析,但对行业指数收益率的分析则较少。为了弥补文献的不足,同时也考虑到影响个股收益的因素多种多样,其随机性一般较大,从行业层面分析因其具有平均效应可使分析结论具有更高的可信度和代表性,所以,本文选择行业指数作为研究对象。

有众多学者致力于证券定价的研究工作。早期的研究中,经济学者倾向于从一般均衡分析出发寻找证券价格的决定机制,具有代表性研究的如 Arrow and Debreu(1954)<sup>[1]</sup>,但证券市场存在高度不确定性,一般均衡分析框架要求对现实进行极大简化,这就注定早期的一般均衡分析不可能很好地解决证券资产定价问题。Markowitz(1952)<sup>[2]</sup>的研究避开一般均衡分析框架,建立了著名的证券组合理论,其最大的贡献在于给出了证券资产的收益与风险的数学定义,这实际上提出了收益是对承受风险的补偿的资产定价理念。后续的研究基本上沿用了Markowitz(1952)的定价思想,其中最受关注的当属 Sharpe

(1964)<sup>[3]</sup>、Lintner(1965)和 Black(1972)提出的资本资产定价模型(CAPM),CAPM表明证券收益的唯一决定因素是该证券与市场组合的 $\beta$ 系数。不过,许多实证研究发现诸如盈余价格、账面市值比等代表公司特征的因素也会影响证券的收益,这是CAPM所无法解释的,这些因素被称为异象。关于异象有很多研究,最具代表性的研究要数 Fama and French(1993),他们发现诸如规模因子和账面市值比等因素可以很好地解释与 $\beta$ 、规模因子、账面市值比、市盈率、财务杠杆等有关的股票的收益。据此他们构建了一个包括超额收益率、规模因子和账面市值比的三因素模型(Fama-French三因素模型,以下简称F-F三因素模型)来解释投资组合的收益率。F-F三因素模型在国外许多实证研究中均得到证实,并因其能很好地解释证券收益而受到广泛的推崇。

近来有很多国内学者关注FF三因素模型在我国证券市场中的应用。关于FF三因素模型的文献,主要有三类:一是检验该模型在我国证券市场的适用性,如范龙振和余世典(2002)<sup>[4]</sup>、杨忻和陈展辉(2003)、邓长荣和马永开(2005)<sup>[5]</sup>以及夏巽(2009)的实证检验表明三因素模型在我国证券市场上是成立

收稿日期 2012-07-29

基金项目:本文为国家社科基金项目“非参数空间计量经济模型及应用研究”(10CTJ002)和中央高校基本科研业务经费项目“动态时变特征的空间计量模型及其在中国地方政府财税策略互动中的应用”(ZK1021)的部分研究成果。

作者简介:方丽婷(1983-),女,福建宁德人,博士研究生,研究方向:统计学;李坤明(1983-),男,福建漳州人,博士研究生,研究方向:理论与应用计量经济学。

的,吴世农和许年行(2004)以及肖军和徐信忠(2004)<sup>[6]</sup>发现三因子模型比CAPM模型更能解释A股股票横截面收益的变化,赵华和吕雯(2010)的研究发现,规模因素、市场因素和账面市值比因素的解能力依次递减;二是以三因素模型为基础进行诸如股票超额联动等实证检验,如蔡伟宏(2006)通过三因素模型对行业指数的收益率进行解释并据此探讨股票市场行业指数超额联动的问题,廖理和沈红波(2008)采用三因子模型研究了中国股权分置改革的市场效应及其影响因素;三是对传统的三因素模型进行改进,如贺炎林(2008)<sup>[7]</sup>、毛小元等(2008)与周芳和张维(2011)的研究。

国内学者在三因素模型的实证应用方面进行了有益探讨,但是,他们的研究通常是使用传统的均值回归方法,然而此方法只是从解释变量平均水平的角度出发考察其对因变量的影响程度,无法基于不同的水平全面地认识两者间的关系,以及不同水平下这种关系的强弱。鉴于FF三因素模型的广泛适用性,同时也为了避免均值回归方法的局限性,我们采用分位数回归(Quantile Regression)方法估计上证行业指数收益率的三因素模型,以全面考察三因素对上证行业指数的影响。

## 二、分位数回归模型简介

分位数回归的思想最早是由Koenker和Bassett于1978年提出的,它是对以古典条件均值模型为基础的最小二乘法的拓展,普通最小二乘法通过使残差平方和最小来估计回归参数,分位数回归则利用因变量的条件分位数来建模,通过最小化加权的残差绝对值之和来估计参数,也即是加权的“最小一乘回归”。与普通的均值回归相比,分位数回归能充分反映自变量对因变量的分布位置、刻度和形状的影响,能够捕捉分布的尾部特征,当自变量对不同部分的因变量的分布产生不同的影响时。例如出现左偏或右偏的情况时,它能更加全面地刻画分布的特征,从而得到全面的分析,而且其分位数回归系数估计比OLS回归系数估计更稳健。

假设 $Y$ 为被解释变量, $X=(X_1, X_2, \dots, X_p)'$ 被解释变量,则分位数回归模型的数学表达式如下:

$$Q_{\tau}(Y|X=x) = x'\beta_{\tau} \quad (1)$$

$Q_{\tau}(Y|X=x)$ 表示在自变量 $X=x$ 的条件下, $Y$ 的第 $\tau$ 分位数。可通过求解如下最优化问题来估计系数 $\beta_{\tau}$ :

$$\hat{\beta}_{\tau} = \arg \min_{\beta} \sum_{i=1}^p \rho_{\tau}(y_i - x_i'\beta) \quad (2)$$

其中 $\rho_{\tau}(z) = \tau z I_{[0, \infty)}(z) + (1 - \tau) \tau z I_{(-\infty, 0)}(z)$ ,  $I(\cdot)$ 为示性函数。

分位数回归模型的参数估计方法主要有单纯形算法(Simplex Method)、内点算法(Interior Point Method)和平滑算法(Smoothing Method)。单纯形法在处理大型数据时运算的速度会显著的降低,但该算法估计出来的参数与其他算法相比具有更高的稳定性和精确度。由于本文的数据不是大型数据,因而拟采用单纯形算法进行参数估计。

## 三、数据与模型设定

### (一)样本数据

本文以上海证券交易所的五个行业指数为研究对象,分别是工业指数、商业指数、地产指数、公用指数、综合指数,基于数据的可获得性,以2010年1月4日至2012年3月6日的日收益率序列为研究数据,在下文中分别记为 $R_{gye}$ 、 $R_{shy}$ 、 $R_{dch}$ 、 $R_{gy}$ 、 $R_{zh}$ ,共525组观测值。数据来源于Wind数据库。

因为上交所行业指数的编制涵盖上交所的A、B股,所以我们用涵盖A、B股的上证综指的日收益率代表市场风险因子,记为 $R_m$ ,Fama and French(1993)构造了能反映账面市值比和小公司风险的组合,用这种组合的收益率代表账面市值比因子和规模因子,后续研究者也采用相同的原理构造这两个因子。本文借鉴蔡伟宏(2006)的做法用上证风格系列指数的日交易数据构造规模因子和账面市值比因子,具体做法如下:

规模因子(SMB)=上证小盘收益率-上证50收益率

账面市值因子(HML)=上证全指成长收益率-上证全指价值收益率

实证中有关收益率的算法通常有两种:一是对数值作差法,即 $R_t = 100(\ln p_t - \ln p_{t-1})$ ,另一种是百分比法,即 $R_t = 100(p_t - p_{t-1})/p_{t-1}$ 。对收益率序列取对数通常可以使序列变得平稳,许多相关研究表明股指的对数收益率往往比百分比收益率更接近正态分布,同时,对数收益率具有可累加性,因而本文采用第一种算法。

### (二)数据的描述性统计和单位根检验

从表1可以看到,各行业指数的平均收益率存在明显差异,除了工业指数,其他四类行业指数的平均收益率均小于市场收益率,而我们所选择的规模因子和账面市值因子的报酬率均为正,这与三因素模型的结论一致。同时,表1还显示,各指数和因子的收益率均表现为不同程度的左偏和尖峰形态,

Jarque-Bera 正态性检验结果表明各指数和因子收益率均不服从正态分布。此外,我们还用两种普遍使用的检验方法对数据进行单位根检验,检验结果说明各数据均是平稳序列。

很显然,由于各数据均不服从正态分布,如果采用普通最小二乘法进行分析,将导致估计结果不可靠,因此有必要使用对数据异常情况更具耐抗性、估计结果更具稳健性的方法,分位数回归方法提供了一种不错的选择。

表1 变量的描述性统计、单位根检验和正态性检验结果

统计量	R_gye	R_shy	R_dch	R_gy	R_zh	Market	SMB	HML
均值	-0.04740	-0.06800	-0.06698	-0.07409	-0.06196	-0.05602	0.03477	0.01232
中位数	0.00202	-0.01064	-0.07122	-0.03655	-0.04046	0.00345	0.15990	0.05775
标准差	1.39453	1.57845	1.74876	1.39684	1.28303	1.30246	1.10944	0.34045
偏度	-0.33974	-0.47818	-0.21642	-0.42414	-0.22822	-0.39394	-0.64745	-0.56161
峰度	4.11022	3.89436	4.32709	4.60344	4.65443	4.42025	4.47478	4.07504
ADF 检验	-22.550	-21.447	-23.741	-22.209	-24.314	-23.296	-17.384	-18.693
PP 检验	-23.732	-21.415	-23.789	-23.283	-24.517	-24.399	-16.041	-18.277
JB 检验	37.062	37.505	42.624	71.982	64.432	57.703	84.257	52.879

注:ADF 检验、PP 检验、JB 检验的统计量均表明在 1%的显著性水平下通过检验。

(三)模型设定

Fama and French(1993)提出的三因素模型如下:

$E(R_i)-R_{ft}=b_1(E(R_M)-R_{ft})+b_2SMB+b_3HML$  (3)

其中  $R_i$ 、 $R_{ft}$  和  $R_M$  分别表示证券收益率、无风险利率和市场收益率,  $E$  表示期望,  $SMB$  和  $HML$  分别是由于公司的规模不同引起的风险溢价和由于公司账面价值比不同引起的溢价。

在进行实证分析的时候,可以根据模型(3)用如下随机模型表述:

$R_{it}-R_{ft}=\alpha+\beta_M(R_{Mt}-R_{ft})+\beta_S SMB_t+\beta_H HML_t+\epsilon_t$  (4)

其中下标  $t$  表示时期,  $\epsilon_t$  为随机扰动项,  $\alpha$ 、 $\beta_M$ 、 $\beta_S$  和  $\beta_H$  为待估系数。

根据本文的研究目的,我们在整个分布的意义上考虑模型(4),而不仅仅是其均值,最终建立如下三因素分位数回归模型:

$Q_\tau(Y|X=x_i)=\alpha_\tau+x_i'\beta_\tau$  (5)

其中  $Y=R_{it}-R_{ft}$ ,  $X=(R_{Mt}-R_{ft}, SMB, HML)'$ ,  $x_i=(R_{Mt}-R_{ft}, SMB_t, HML_t)'$ ,  $\beta_\tau=(\beta_{M\tau}, \beta_{S\tau}, \beta_{H\tau})$ 。

四、实证结果及分析

我们接着用本文第二部分所介绍的分位数估计方法对模型(5)进行估计,主要分位点的估计结果见表2,整体分位点的系数估计结果见图1。为了进行比较,我们还用普通最小二乘法估计了模型(结果见

表3)。

表2 主要分位点的估计结果

分位点	变量	R_gye	R_shy	R_dch	R_gy	R_zh
0.1	截距	-0.2425***	-0.6579***	-1.0932***	-0.4379***	-0.3352***
	RM	1.0280***	1.0163***	1.0851***	1.0116***	0.9293***
	SMB	0.1408***	0.4986***	-0.2003***	0.3371***	-0.2319***
	HML	0.1913***	0.3092***	0.1560***	-0.5399***	-0.1563***
0.2	截距	-0.1402***	-0.4084***	-0.7186***	-0.3103***	-0.2183***
	RM	1.0352***	1.0094***	1.0653***	1.0033***	0.9404***
	SMB	0.1114***	0.5048***	-0.2082***	0.3150***	-0.2337***
	HML	0.2536***	0.2709***	0.4998***	-0.5397***	-0.1630***
0.3	截距	-0.0811***	-0.2644***	-0.4120***	-0.2039***	-0.1255***
	RM	1.0385***	1.0145***	1.1224***	0.9920***	0.9376***
	SMB	0.1201***	0.5177***	-0.1841***	0.3193***	-0.2459***
	HML	0.2453***	0.2484***	0.2557***	-0.5513***	-0.1982***
0.4	截距	-0.0406***	-0.1558***	-0.2374***	-0.1354***	-0.0558***
	RM	1.0398***	1.0192***	1.1409***	1.0009***	0.9384***
	SMB	0.1167***	0.5312***	-0.1735***	0.3131***	-0.2306***
	HML	0.2635***	0.1283***	0.2561***	-0.5879***	-0.2961***
0.5	截距	0.0036***	-0.0624***	-0.0699***	-0.0455***	0.0035***
	RM	1.0313***	1.0123***	1.0976***	1.0077***	0.9453***
	SMB	0.1145***	0.5417***	-0.2043***	0.3123***	-0.2364***
	HML	0.2773***	0.1407***	0.3349***	-0.6148***	-0.3048***
0.6	截距	0.0502***	0.0757***	0.1235***	0.0483***	0.0579***
	RM	1.0297***	1.0132***	1.1118***	1.0045***	0.9501***
	SMB	0.0972***	0.5485***	-0.2301***	0.3080***	-0.2511***
	HML	0.2810***	0.1396***	0.3367***	-0.5895***	-0.2901***
0.7	截距	0.0901***	0.2140***	0.3764***	0.1508***	0.1127***
	RM	1.0392***	0.9908***	1.1130***	1.0002***	0.9437***
	SMB	0.0868***	0.5118***	-0.2876***	0.3172***	-0.2504***
	HML	0.2837***	0.1912***	0.5271***	-0.6347***	-0.3112***
0.8	截距	0.1555***	0.3730***	0.7374***	0.2731***	0.2122***
	RM	1.0460***	0.9893***	1.1383***	0.9939***	0.9455***
	SMB	0.0896***	0.4561***	-0.2196***	0.3763***	-0.2654***
	HML	0.2567***	0.3459***	0.0045***	-0.7921***	-0.2252***
0.9	截距	0.2500***	0.5672***	1.1835***	0.4439***	0.3364***
	RM	1.0405***	1.0077***	1.1656***	0.9781***	0.9590***
	SMB	0.1092***	0.4706***	-0.2313***	0.3731***	-0.2489***
	HML	0.2056***	0.2878***	0.0212***	-0.9184***	-0.3251***

注:\*\*\*、\*\*、\*分别表示在 1%、5%和 10%的水平下显著。

观察表2中分位数回归估计结果,我们发现在绝大多数的分位点上,各个行业指数的影响因子的系数估计值在 1%或者 5%显著性水平上均显著。尤其是工业指数、公用指数、综合指数的影响因子在本文所考察的所有分位点上的系数估计值均显著。对商业指数和地产指数而言,账面市值因子在部分分位点上的系数估计值不显著或显著性水平不高,而其他两个因子在所有分位点上的系数估计值均显著。此外,五个行业指数的市场因子在所有分位点上的系数估计值均为正且数值较稳定,波动较小,说明市场收益率对各个行业指数收益率具有稳定的正向影响。规模因子对五个行业指数的影响方向不同,其中对综合指数和地产指数具有负向影响,而对其他三类指数具有正向影响,其系数估计值也相对平稳,但是波动性比市场因子大。账面市值因子对五个行业指数的影响方向也不同,其中对综合指数和公用指数具有负向影响,对工业指数和商业指数



具有正向影响,对地产指数既有正向影响又有负向影响,其中部分系数估计值也相对不平稳,波动性较大。具体而言,账面市值因子对工业指数、地产指数、综合指数的影响程度在中间分位点上(0.4,0.5,0.6,0.7)较平稳,在两端分位点上波动较大。

观察表3中的最小二乘(OLS)估计结果,我们发现五个指数的各个影响因子的系数估计值均显著,其中市场因子的系数估计值均为正,规模因子对综合指数和地产指数具有负向影响,而对其他三类指数具有正向影响,账面市值因子对综合指数和公用指数具有负向影响,对工业指数、商业指数和地产指数具有正向影响。这些估计结果与分位数估计结果基本一致,但是前者仅是后者的平均,无法从各个水平上全面描述各个影响因子对各类指数的影响程度以及影响的强弱。

表3 均值估计结果

自变量	R_gye	R_shy	R_dch	R_gy	R_zh
截距	0.00410 (0.00979)	-0.03310 (0.02228)	0.00032155 (0.04269)	-0.02153 (0.01605)	0.00215 (0.01339)
R <sub>M</sub>	1.03757*** (0.00753)	0.99933*** (0.01715)	1.11603*** (0.03286)	1.00239*** (0.01236)	0.94127*** (0.01031)
SMB	0.10155*** (0.01281)	0.52326*** (0.02916)	-0.25650*** (0.05589)	0.32479*** (0.02101)	-0.24576*** (0.01753)
HML	0.25090*** (0.04167)	0.23403*** (0.09483)	0.33505*** (0.18175)	-0.62564*** (0.06834)	-0.23053*** (0.05702)

注: \*、\*\*、\*\*\* 分别表示统计量在 1%、5% 和 10% 水平下显著。括号中的数值为标准差。

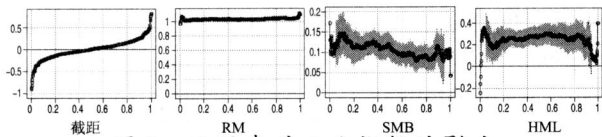


图1 三因素对工业指数的影响

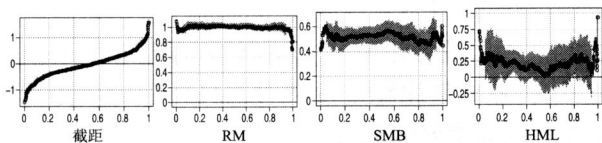


图2 三因素对商业指数的影响

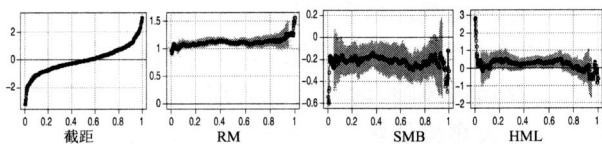


图3 三因素对地产指数的影响

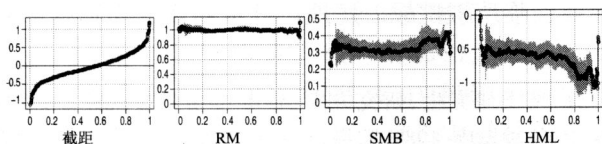


图4 三因素对公用指数的影响

四、结论与启示

本文利用分位数回归方法来估计Fama French三因素模型,进而对上证行业指数收益率的影响因素进行分析考察。分位数回归的估计方法使我们得以从不同的水平完整地描述了各个因子对收益率的作用关系,其效果远远超越传统的回归方法的平均水平对两者关系的描述。实证结果表明,市场因子对五类行业指数均有显著的稳定的正向影响,规模因子对五个行业指数的影响也显著但影响方向不同,其中对综合指数和地产指数具有负向影响,而对其他三类指数具有正向影响,账面市值因子对五个行业指数的影响方向也不同,其中对综合指数和公用指数具有负向影响,对工业指数和商业指数具有正向影响,对地产指数既有正向影响又有负向影响,同时在部分分位点上,账面市值因子对五个行业指数的影响不显著。

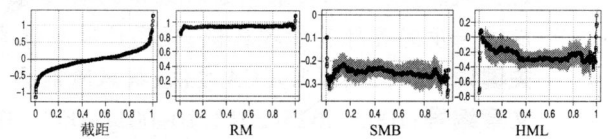


图5 三因素对综合指数的影响

最后,对比普通最小二乘估计结果与分位数估计结果,发现两者基本一致,但是前者仅是后者的平均,无法从各个水平上全面描述各个影响因子对各类指数的影响程度以及影响的强弱,这充分体现了分位数回归估计方法相比于传统的线性回归估计方法的优越性。

参考文献:

[1]Arrow K and Debreu G. Existence of an Equilibrium for a Competitive Economy[J]. Econometrica, 1954, 22(3): 265-290.  
[2]Markowitz M. Portfolio Selection[J]. Journal of Finance, 1952, 7(1): 77-91.  
[3]Sharpe F. Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk[J]. The Journal of Finance, 1964, 19(3): 425-442.  
[4]范龙振,余世典.中国股票市场的三因素模型[J].系统工程学报,2002,17(6):537-546.  
[5]邓长荣,马永开.三因素模型在中国证券市场的实证研究[J].管理学报,2005,2(5):59-64.  
[6]肖军,徐信忠.中国股市价值反转投资策略有效性实证研究[J].经济研究,2004,(3):55-64.  
[7]贺炎林.基于状态转移信息对FF三因素模型的改进[J].中国管理科学,2008,16(1):7-15.

(责任编辑 张艳峰)